LENDING CLUB
-
CASE STUDY

CREATED BY - SAGAR MARU

PLANNING THE PROCESS

Data Understanding: Gain insights into the dataset.

Data Cleaning and Manipulation: Prepare and refine the data for analysis.

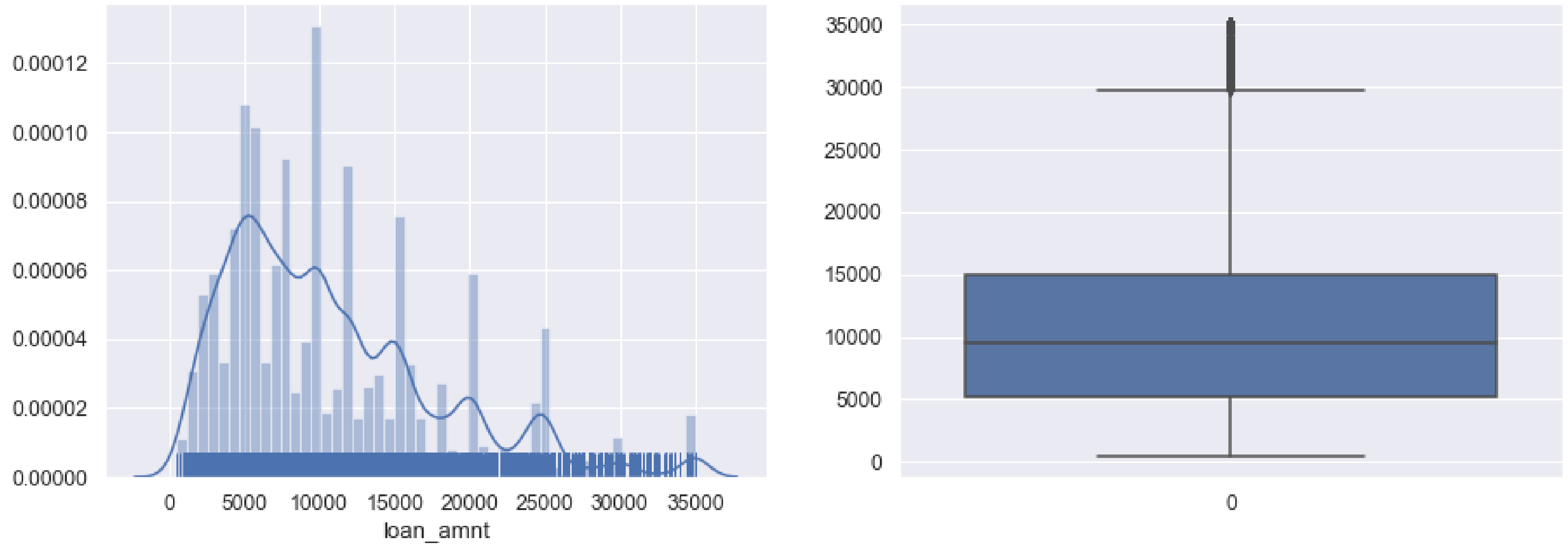
Data Analysis and Visualization: Analyze the data and create visual representations to uncover insights

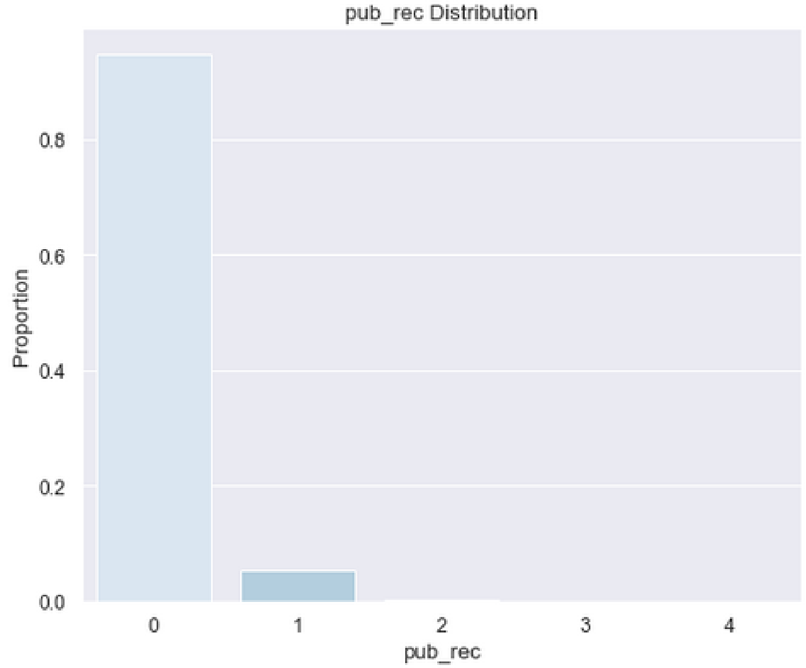
DATA UNDERSTANDING

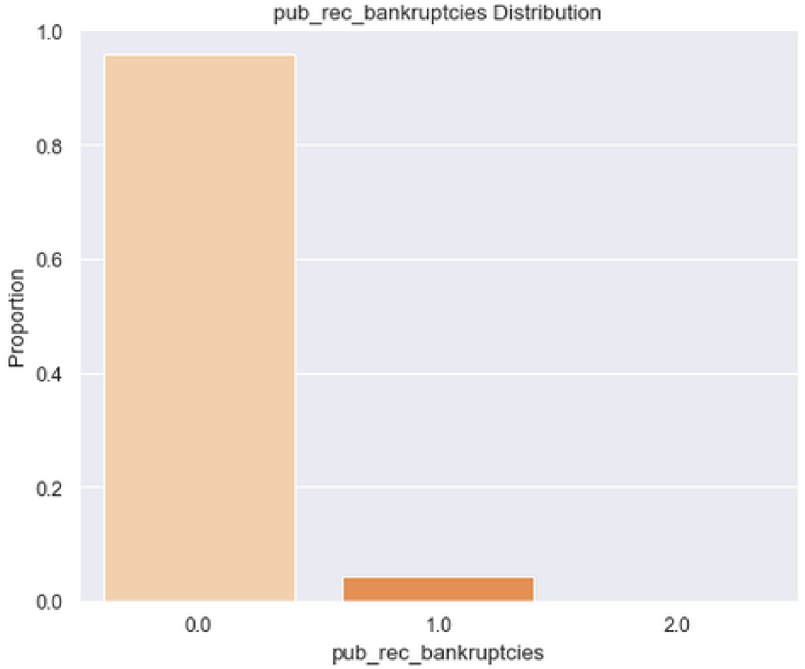
DATA CLEANING AND MANIPULATION

DATA ANALYSIS AND VISUALIZATION

Loan Amount Distribution Chart







Loan Amount:

A substantial number of people took out loans of 9,500, which aligns with the median of the distribution. Loans exceeding 30,000 are relatively rare.

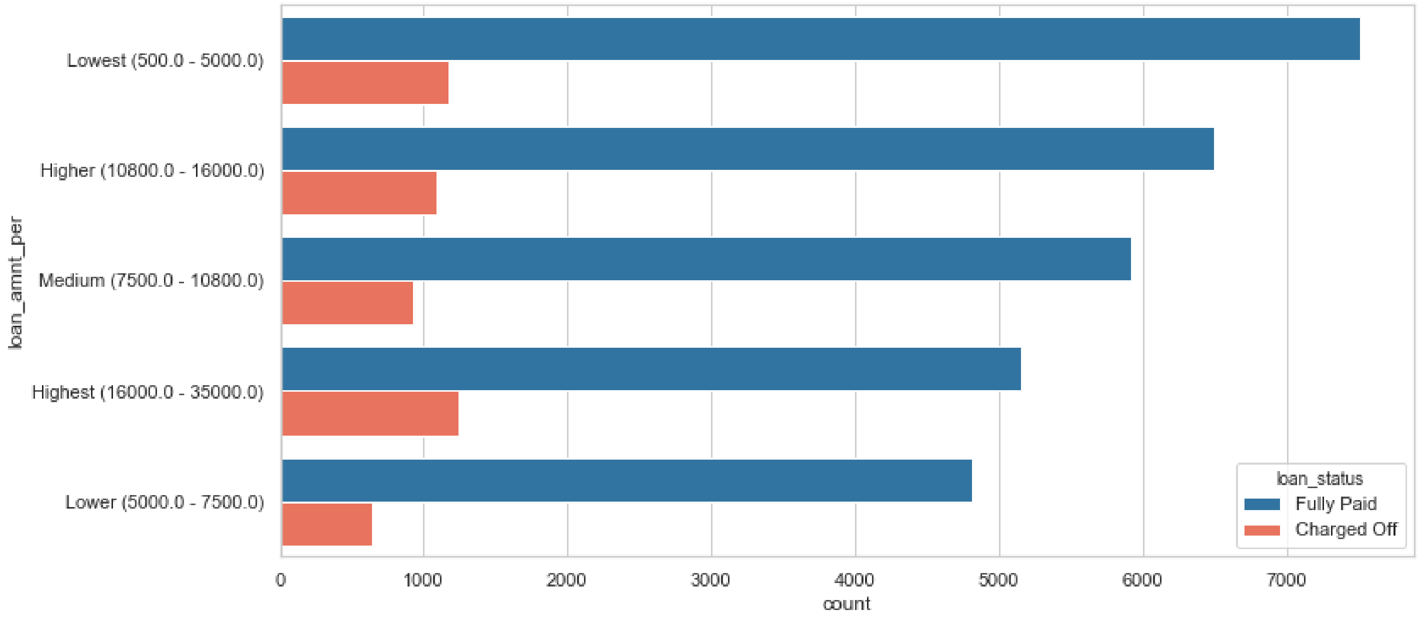
Public Records:

Additionally, approximately 90% of borrowers have no public derogatory records, and 99% have not declared bankruptcy.

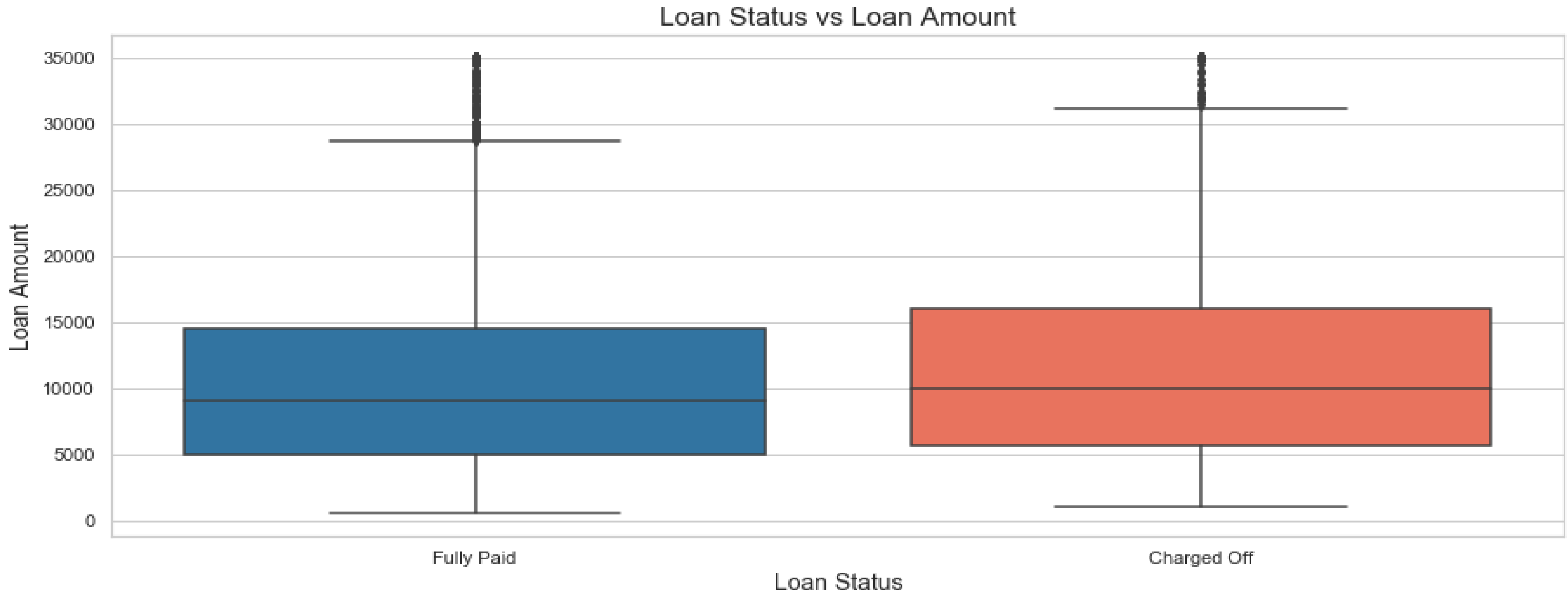
INTEREST RATE VS LOAN STATUS

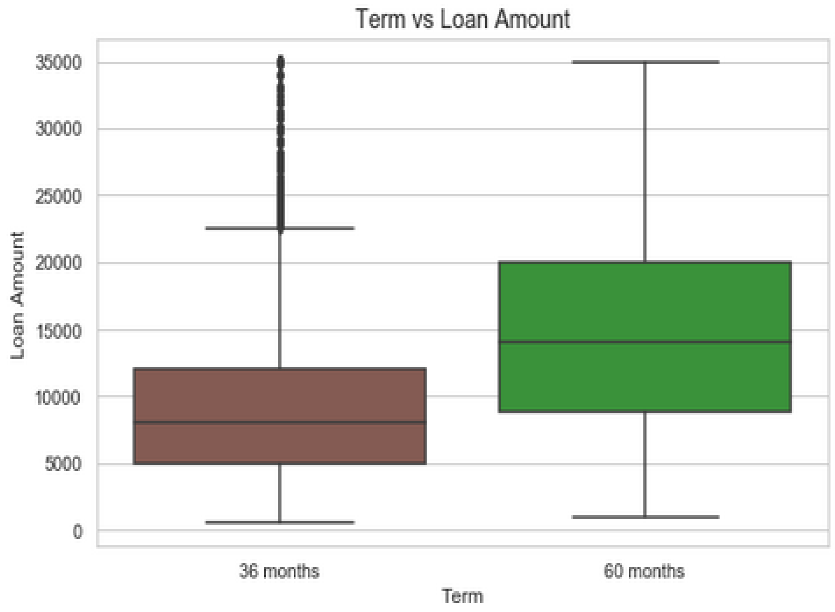
Larger loan amounts are more likely to result in default.

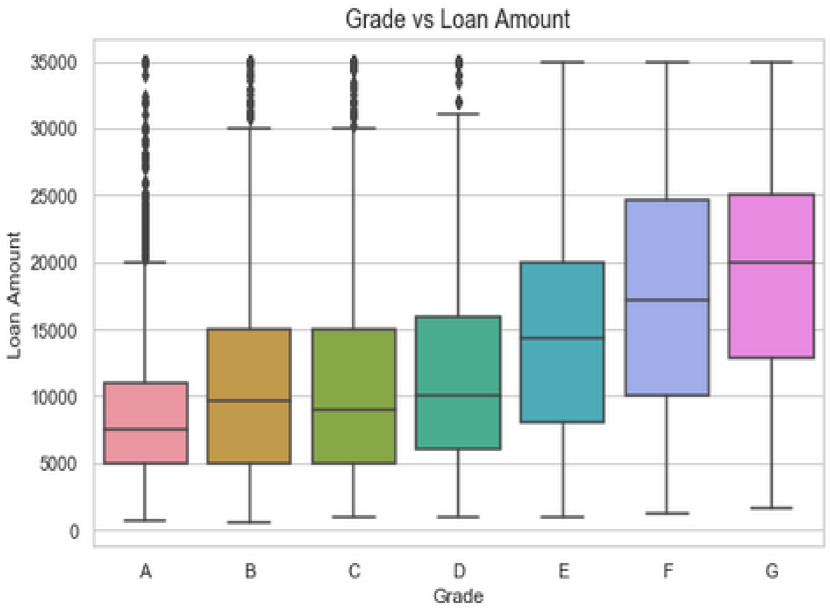
This trend suggests that borrowers with higher loan balances face a greater risk of non-repayment, indicating a direct correlation between the loan size and the likelihood of default.



loan_amnt_per	Charged off %	Record count
Highest (16000.0 - 35000.0)	0.193649	6393
Higher (10800.0 - 16000.0)	0.143309	7585
Medium (7500.0 - 10800.0)	0.135163	6851
Lowest (500.0 - 5000.0)	0.134998	8689
Lower (5000.0 - 7500.0)	0.118262	5454





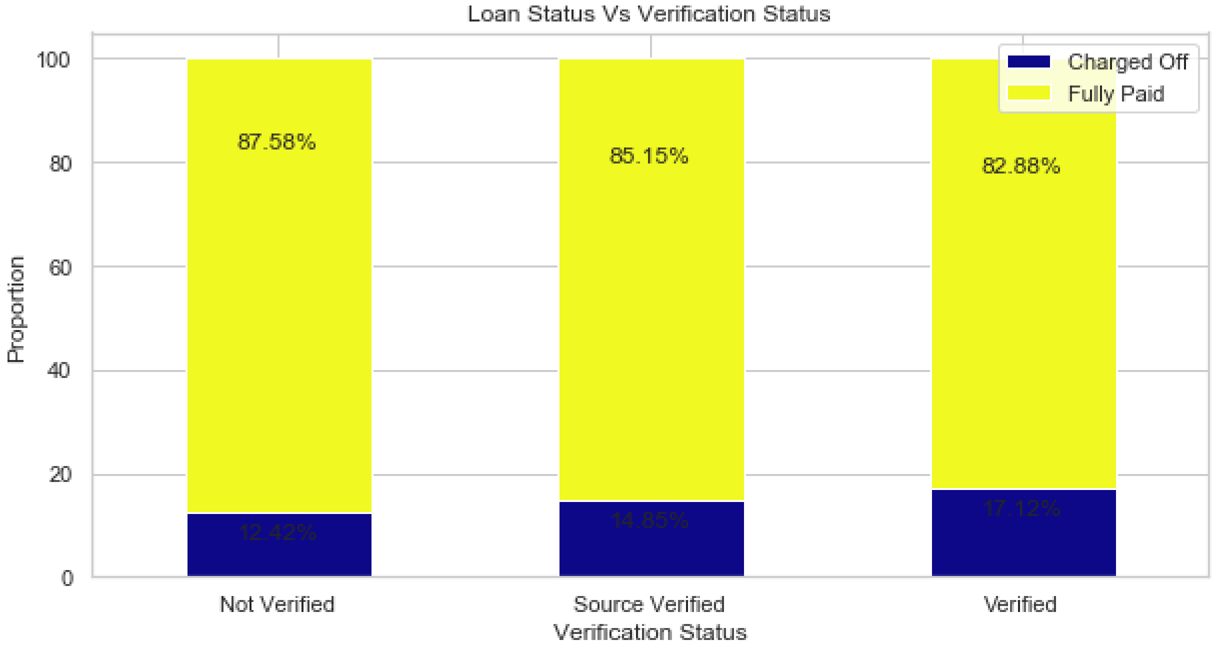


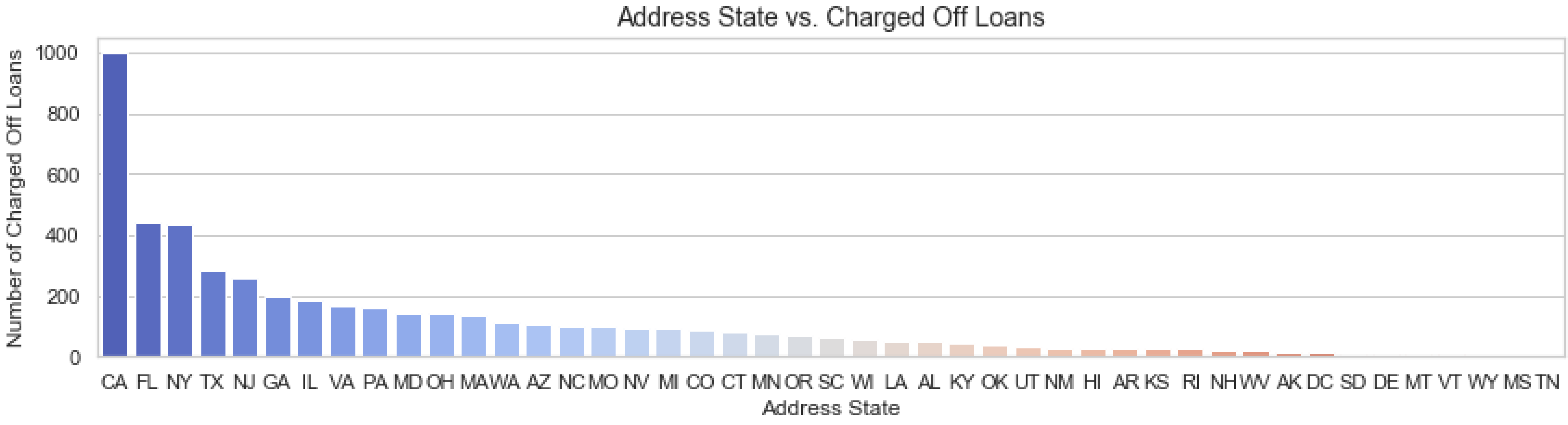
Charged Off loans typically have higher amounts than Fully Paid ones, suggesting larger loans are more prone to default.

Higher loan amounts often correspond with longer tenures, especially 60 months. Borrowers with grades 'F' and 'G' tend to take the largest loans, with amounts increasing from grade A to G.

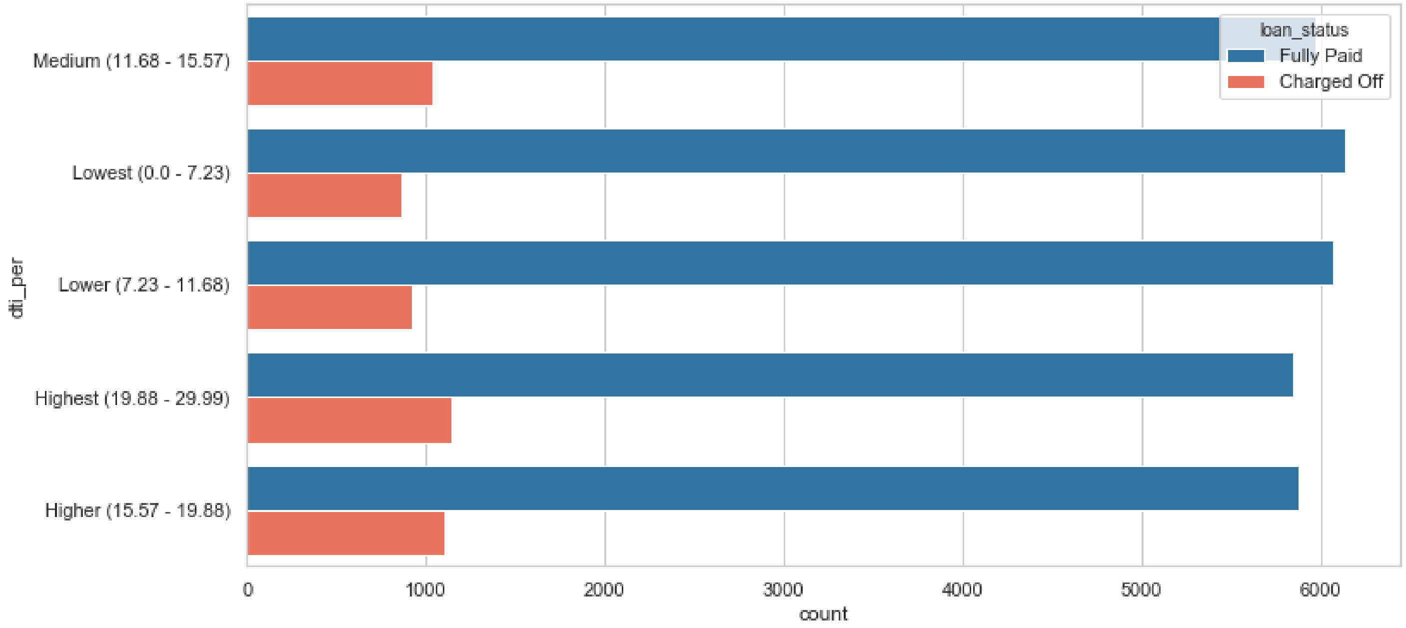
<u>Verification Status:</u>
Interestingly, verified applicants have a higher default rate, indicating potential flaws in the verification process. Surprisingly, borrowers with unverified incomes are more likely to repay fully.

<u>Address State:</u>
The highest defaults occurred in CA, FL, and NY, followed by TX and NJ, highlighting regional differences in loan performance and higher risk in these states.





DEBT-TO-INCOME (DTI) RATIO VS LOAN STATUS



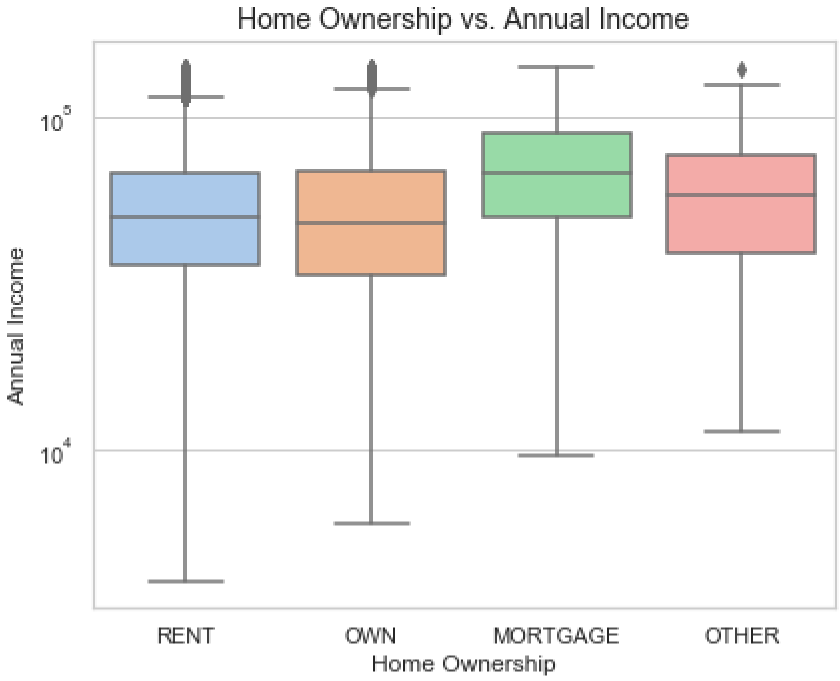
dti_per	Charged off %	Record count
Highest (19.88 - 29.99)	0.163709	6988
Higher (15.57 - 19.88)	0.158219	6984
Medium (11.68 - 15.57)	0.147361	7010
Lower (7.23 - 11.68)	0.132046	6990
Lowest (0.0 - 7.23)	0.123429	7000

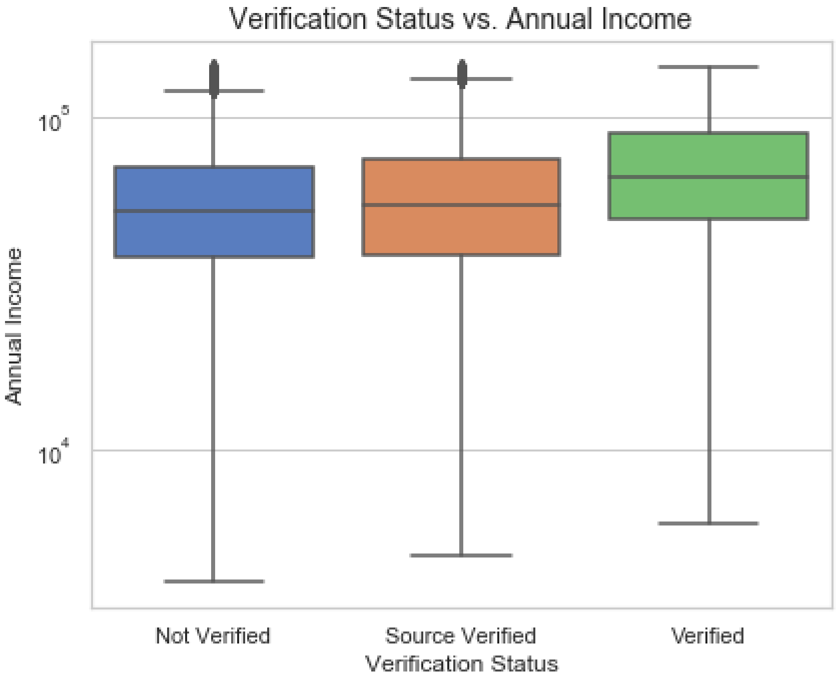
A higher Debt-to-Income (DTI) ratio is linked to a higher percentage of charged-off loans.

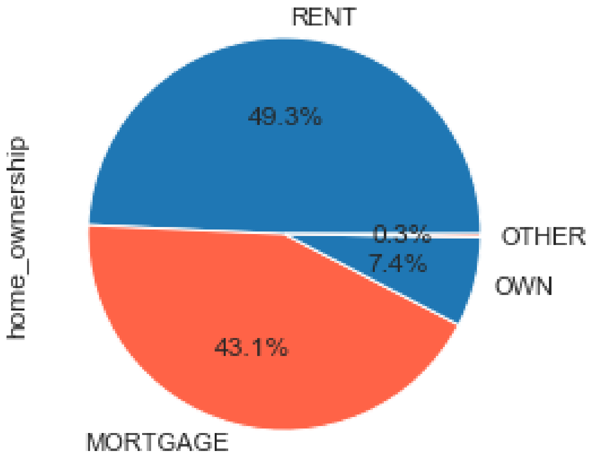
As borrowers accumulate more debt relative to their income, the risk of default increases, highlighting DTI as a crucial factor in evaluating loan repayment capability.

Individuals with mortgages typically have higher annual incomes compared to other home ownership statuses.

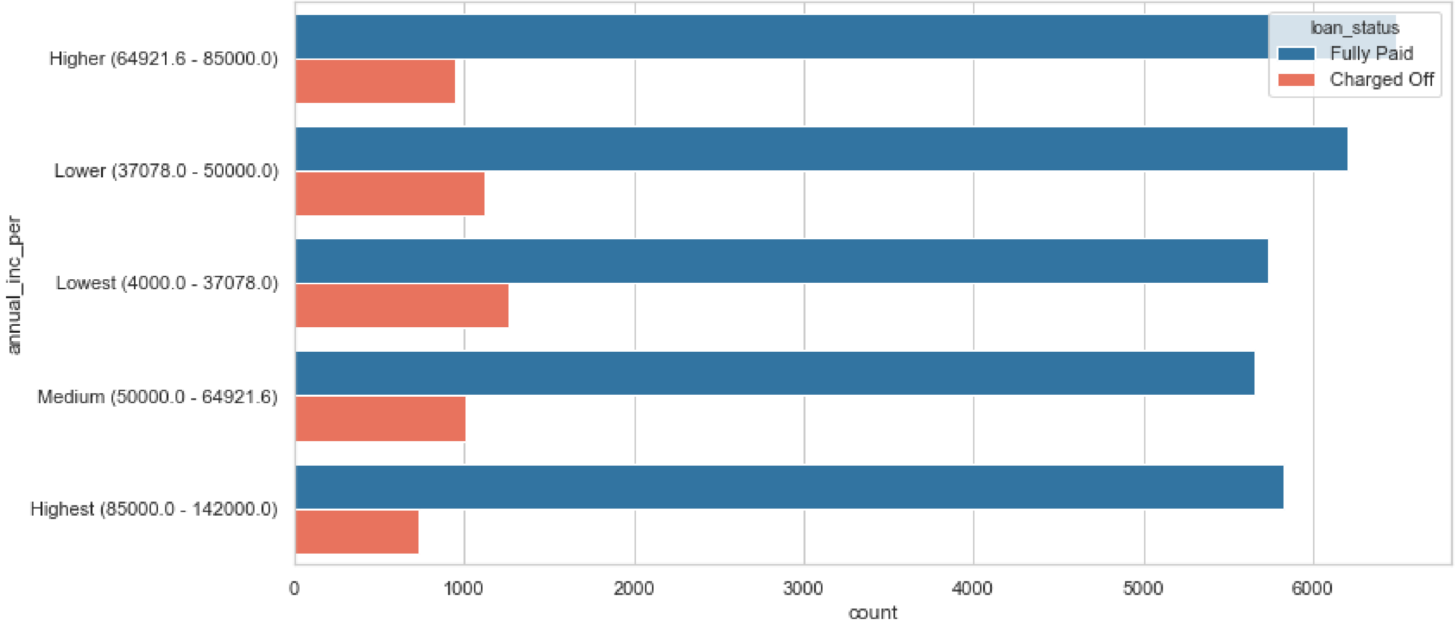
Higher-income borrowers are also more likely to have verified income sources, indicating a strong link between mortgage ownership, income verification, and income levels. These insights highlight the relationship between home ownership status and financial verification.







ANNUAL INCOME VS LOAN STATUS



annual_inc_per	Charged off %	Record count
Lowest (4000.0 - 37078.0)	0.179700	6995
Lower (37078.0 - 50000.0)	0.152996	7327
Medium (50000.0 - 64921.6)	0.151479	6661
Higher (64921.6 - 85000.0)	0.127926	7434
Highest (85000.0 - 142000.0)	0.111518	6555

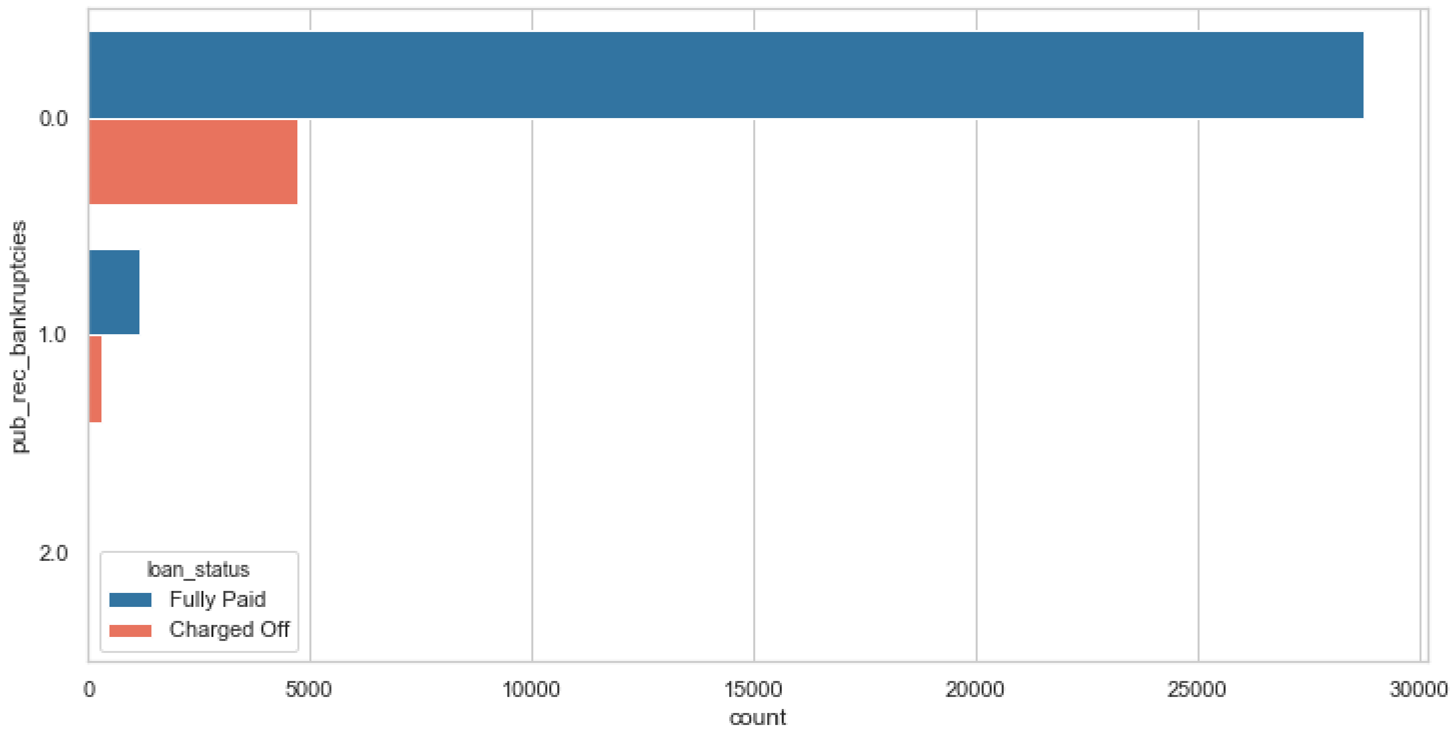
Higher income levels are linked to better repayment rates, making defaults less likely.

Conversely, lower income levels increase the chances of loan defaults. This highlights the significant role income plays in repayment behavior, with higher income reducing the risk of default.

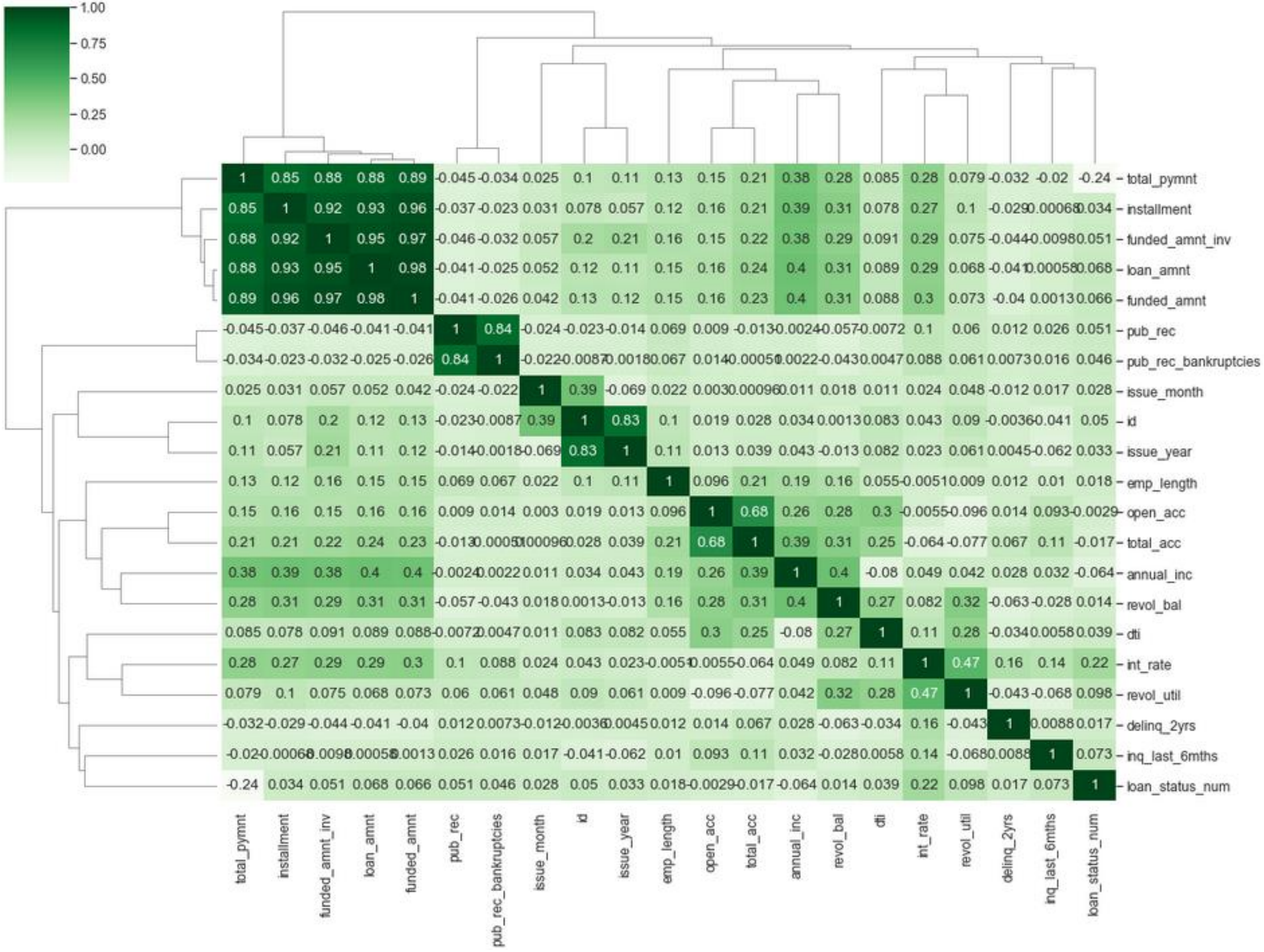
PUBLIC BANKRUPTCY RECORDS VS LOAN STATUS

A higher number of public bankruptcy records significantly increases the risk of loan default.

More records suggest a greater likelihood of non-repayment, making this an important factor in evaluating loan risk.



pub_rec_bankruptcies	Charged off %	Record count
2.0	0.400000	5
1.0	0.220884	1494
0.0	0.141517	33473



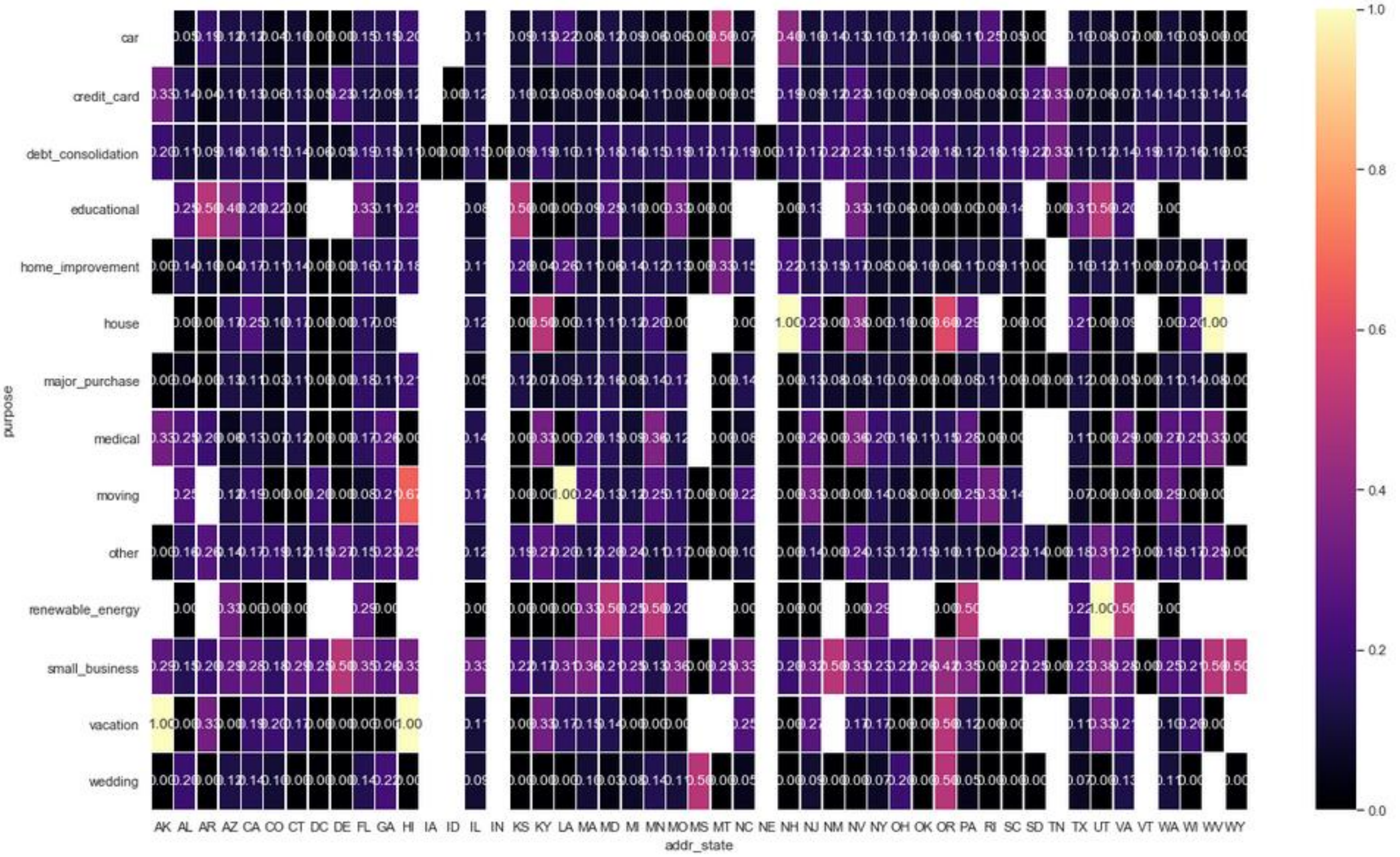
Darker values indicate stronger correlations. Loan_amnt, funded_amnt, funded_amnt_inv, installment, and total_pymnt are highly correlated, showing strong relationships.

Additionally, public records (pub_rec, pub_rec_bankruptcies) and account-related fields (open_acc, total_acc) exhibit significant correlation, indicating these variables are closely connected and may influence each other within the dataset.

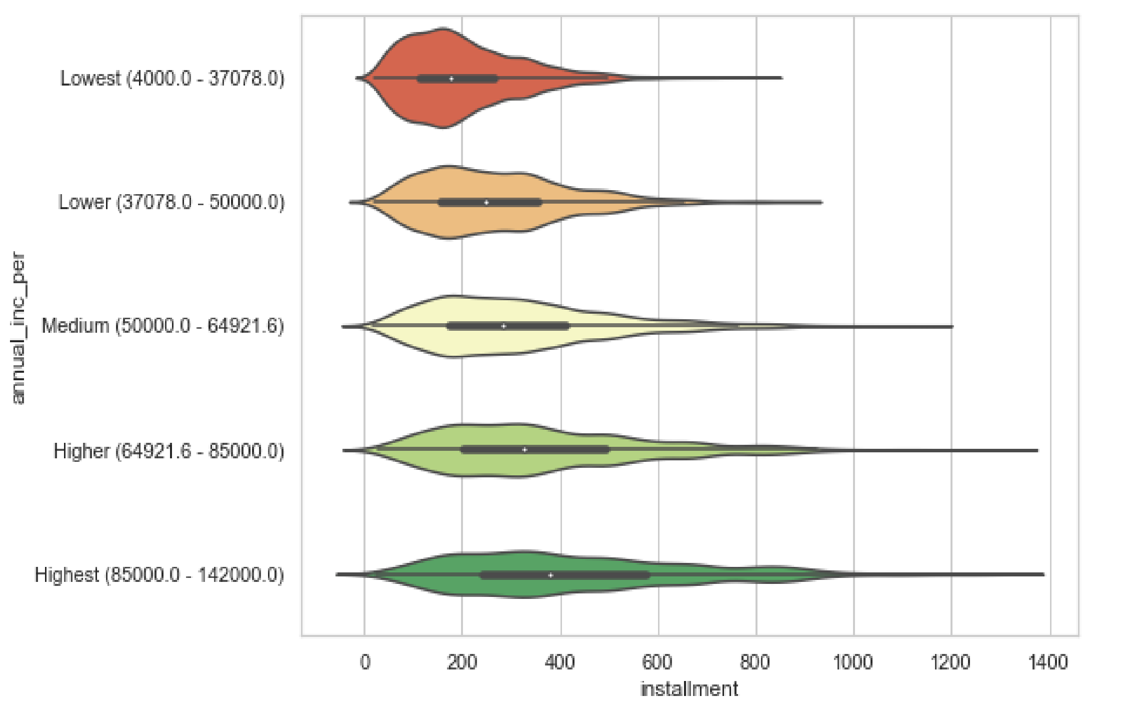
Vacation Loans: Higher risk is observed in states such as Alaska (AK), Hawaii (HI), and Oregon (OR).

Education Loans: Elevated risk is noted in Arkansas (AR), Kansas (KS), and Utah (UT).

Small Business Loans: Delaware (DE), New Mexico (NM), West Virginia (WV), and Wyoming (WY) show higher risks for small business loans.
Wedding Loans: Increased risk is observed in Mississippi (MS) and Oregon (OR).



Higher installment amounts are linked to more defaults across all income groups, indicating that as payments increase, the likelihood of loan defaults rises, regardless of income level.



THANK YOU

From – Sagar Maru